EXPECTED NUMBER OF REAL ZEROS FOR RANDOM LINEAR COMBINATIONS OF ORTHOGONAL POLYNOMIALS

D. S. LUBINSKY, I. E. PRITSKER, AND X. XIE

(Communicated by Walter Van Assche)

ABSTRACT. We study the expected number of real zeros for random linear combinations of orthogonal polynomials. It is well known that Kac polynomials, spanned by monomials with i.i.d. Gaussian coefficients, have only $(2/\pi + o(1)) \log n$ expected real zeros in terms of the degree n. On the other hand, if the basis is given by Legendre (or more generally by Jacobi) polynomials, then random linear combinations have $n/\sqrt{3} + o(n)$ expected real zeros. We prove that the latter asymptotic relation holds universally for a large class of random orthogonal polynomials on the real line, and also give more general local results on the expected number of real zeros.

1. BACKGROUND

Zeros of polynomials with random coefficients have been intensively studied since 1930s. The early work concentrated on the expected number of real zeros $\mathbb{E}[N_n(\mathbb{R})]$ for polynomials of the form $P_n(x) = \sum_{k=0}^{n} c_k x^k$, where $\{c_k\}_{k=0}^{n}$ are independent and identically distributed random variables. Apparently the first paper that initiated the study is due to Bloch and Pólya [2]. They gave an upper bound $\mathbb{E}[N_n(\mathbb{R})] = O(\sqrt{n})$ for polynomials with coefficients selected from the set $\{-1, 0, 1\}$ with equal probabilities. Further results generalizing and improving that estimate were obtained by Littlewood and Offord [21]-[22], Erdős and Offord [8] and others. Kac [17] established the important asymptotic result

$$\mathbb{E}[N_n(\mathbb{R})] = (2/\pi + o(1)) \log n \quad \text{as } n \to \infty,$$

for polynomials with independent real Gaussian coefficients. More precise forms of this asymptotic were obtained by many authors, including Kac [18], Wang [32], Edelman and Kostlan [7]. It appears that the sharpest known version is given by the asymptotic series of Wilkins [33]. Many additional references and further directions of work on the expected number of real zeros may be found in the books of Bharucha-Reid and Sambandham [1], and of Farahmand [9]. In fact, Kac [17]-[18] found the exact formula for $\mathbb{E}[N_n(\mathbb{R})]$ in the case of standard real Gaussian coefficients:

$$\mathbb{E}[N_n(\mathbb{R})] = \frac{4}{\pi} \int_0^1 \frac{\sqrt{A(x)C(x) - B^2(x)}}{A(x)} \, dx,$$

where

$$A(x) = \sum_{j=0}^{n} x^{2j}, \quad B(x) = \sum_{j=1}^{n} jx^{2j-1} \quad \text{and} \quad C(x) = \sum_{j=1}^{n} j^2 x^{2j-2}.$$

In the subsequent paper Kac [19], the asymptotic result for the number of real zeros was extended to the case of uniformly distributed coefficients on $[-1, 1]$. Erdős and Offord [8] generalized the Kac asymptotic to Bernoulli distribution (uniform on $\{-1, 1\}$), while Stevens [28] considered a wide class of distributions. Finally, Ibragimov and Maslova [15, 16] extended the result to all mean-zero distributions in the domain of attraction of the normal law.

We state a result on the number of real zeros for the random linear combinations of rather general functions. It originated in the papers of Kac [17]-[19], who used the monomial basis, and was extended to trigonometric polynomials and other bases, see Farahmand [9] and Das [4]-[5]. We are particularly interested in the bases of orthonormal polynomials, which is the case considered by Das [4]. For any set $E \subset \mathbb{C}$, we use the notation $N_n(E)$ for the number of zeros of random functions (1.1) (or random orthogonal polynomials of degree at most n) located in E. The expected number of zeros in E is denoted by $\mathbb{E}[N_n(E)]$, with $\mathbb{E}[N_n(a, b)]$ being the expected number of zeros in $(a, b) \subset \mathbb{R}$.

Proposition 1.1. *Let $[a, b] \subset \mathbb{R}$, and consider real valued functions $g_j(x) \in C^1([a, b])$, $j = 0, \ldots, n$, with $g_0(x)$ being a nonzero constant. Define the random function*

$$(1.1) \qquad\qquad G_n(x) = \sum_{j=0}^{n} c_j g_j(x),$$

where the coefficients c_j are i.i.d. random variables with Gaussian distribution $\mathcal{N}(0, \sigma^2)$, $\sigma > 0$. If there is $M \in \mathbb{N}$ such that $G_n'(x)$ has at most M zeros in (a, b) for all choices of coefficients, then the expected number of real zeros of $G_n(x)$ in the interval (a, b) is given by

$$(1.2) \qquad\qquad \mathbb{E}[N_n(a, b)] = \frac{1}{\pi} \int_a^b \frac{\sqrt{A(x)C(x) - B^2(x)}}{A(x)} \, dx$$

where

$$(1.3) \qquad A(x) = \sum_{j=0}^{n} g_j^2(x), \quad B(x) = \sum_{j=1}^{n} g_j(x)g_j'(x) \quad \text{and} \quad C(x) = \sum_{j=1}^{n} [g_j'(x)]^2.$$

Clearly, the original formula of Kac follows from this proposition for $g_j(x) = x^j$, $j = 0, 1, \ldots, n$. We sketch a proof of Proposition 1.1 in Section 3, as we could not find a suitable reference with a complete proof for Proposition 1.1 in this general form. We note that multiple zeros are counted only once by the standard convention in all of the above results on real zeros. However, the probability of having a multiple zero for a polynomial with Gaussian coefficients is equal to 0, so that we have the same result on the expected number of zeros regardless whether they are counted with or without multiplicities.

2. Random orthogonal polynomials

Let μ denote a positive Borel measure compactly supported on the real line, with infinitely many points in its support, and with finite power moments of all orders. For $n \geq 0$, let

$$p_n(x) = \gamma_n x^n + \ldots$$

denote the nth orthonormal polynomial for μ, with $\gamma_n > 0$, so that

$$\int p_n p_m d\mu = \delta_{mn}.$$

Using the orthonormal polynomials $\{p_j\}_{j=0}^\infty$ as the basis, we consider the ensemble of random polynomials of the form

$$(2.1) \qquad P_n(x) = \sum_{j=0}^n c_j p_j(x), \quad n \in \mathbb{N},$$

where the coefficients $c_0, c_1, \ldots, c_n$ are i.i.d. random variables. Such a family is often called random orthogonal polynomials. If the coefficients have Gaussian distribution, one can apply Proposition 1.1 to study the expected number of real zeros of random orthogonal polynomials. In particular, Das [4] considered random Legendre polynomials, and found that $\mathbb{E}[N_n(-1,1)]$ is asymptotically equal to $n/\sqrt{3}$. Wilkins [34] improved the error term in this asymptotic relation by showing that $\mathbb{E}[N_n(-1,1)] = n/\sqrt{3} + o(n^\varepsilon)$ for any $\varepsilon > 0$. For random Jacobi polynomials, Das and Bhatt [6] concluded that $\mathbb{E}[N_n(-1,1)]$ is asymptotically equal to $n/\sqrt{3}$ too. They also provided estimates for the expected number of real zeros of random Hermite and Laguerre polynomials, but those arguments contain significant gaps. Farahmand [9, 10, 11] considered various generalizations of these results for the level crossings of random sums of Legendre polynomials with coefficients that may have different distributions. Interesting computations and pictures of zeros of random orthogonal polynomials may be found on the CHEBFUN web page of Trefethen [31].

For the orthonormal polynomials $\{p_j(x)\}_{j=0}^\infty$ associated with the measure μ, define the reproducing kernel by

$$K_n(x,y) = \sum_{j=0}^{n-1} p_j(x)p_j(y),$$

and the differentiated kernels by

$$K_n^{(k,l)}(x,y) = \sum_{j=0}^{n-1} p_j^{(k)}(x)p_j^{(l)}(y), \quad k, l \in \mathbb{N} \cup \{0\}.$$

The strategy is to apply Proposition 1.1 with $g_j = p_j$, so that

$$(2.2) \qquad A(x) = K_{n+1}(x,x), \quad B(x) = K_{n+1}^{(0,1)}(x,x) \quad \text{and} \quad C(x) = K_{n+1}^{(1,1)}(x,x).$$

We use universality limits for the reproducing kernels of orthogonal polynomials (see Lubinsky [23]-[24] and Totik [29]-[30]), and asymptotic results on zeros of random polynomials (cf. Pritsker [25]) to give asymptotics for the expected number of real zeros for a wide class of random orthogonal polynomials.

Theorem 2.1. *Let $K \subset \mathbb{R}$ be a finite union of closed and bounded intervals, and let μ be a positive Borel measure supported on K such that $d\mu(x) = w(x)dx$ and $w > 0$ a.e. on K. If for every $\varepsilon > 0$ there is a closed set $S \subset K$ of Lebesgue measure $|S| < \varepsilon$, and a constant $C > 1$ such that $C^{-1} < w < C$ a.e. on $K \setminus S$, then the expected number of real zeros of random orthogonal polynomials (2.1) with Gaussian coefficients satisfy*

$$(2.3) \qquad \lim_{n\to\infty} \frac{1}{n}\mathbb{E}[N_n(\mathbb{R})] = \frac{1}{\sqrt{3}}.$$

3

A simple example of the orthogonality measure μ satisfying the above conditions is given by the density w that is continuous on K except for finitely many points, and has finitely many zeros on K. More specifically, one may consider the generalized Jacobi weight of the form $w(x) = v(x) \prod_{j=1}^{J} |x - x_j|^{\alpha_j}$, where $v(x) > 0$, $x \in K$, and $\alpha_j > -1$, $j = 1, \ldots, J$.

Theorem 2.1 is a consequence of more precise and general local results given below. In order to state them, we need the notion of the equilibrium measure ν_K of a compact set $K \subset \mathbb{C}$. This is the unique probability measure supported on K that minimizes the energy

$$I[\nu] = -\iint \log |z - t| \, d\nu(t) d\nu(z)$$

amongst all probability measures ν with support on K. The logarithmic capacity of K is

$$\mathrm{cap}(K) = \exp\left(-I[\nu_K]\right).$$

When we say that a compact set K is regular, this means regularity in the sense of Dirichlet problem (or potential theory). See Ransford [26] for further orientation.

We also need the notion of a measure μ regular in the sense of Stahl, Totik, and Ullman [27]. If $K = \mathrm{supp}\,\mu$ and

$$\lim_{n \to \infty} \gamma_n^{1/n} = \frac{1}{\mathrm{cap}(K)},$$

where γ_n is the leading coefficient of p_n, then we say that μ is STU-regular. A sufficient condition for this is that K consists of finitely many intervals and $\mu' = w > 0$ a.e. in those intervals.

Theorem 2.2. *Let μ be an STU regular measure with compact support $K \subset \mathbb{R}$, which is regular in the sense of potential theory. Let O be an open set in which μ is absolutely continuous, and such that for some $C > 1$*

$$(2.4) \qquad\qquad C^{-1} \leq \mu' \leq C \text{ a.e. in } O.$$

Then given any compact subinterval $[a, b]$ of O, we have

$$(2.5) \qquad\qquad \lim_{n \to \infty} \frac{1}{n} \mathbb{E}\left[N_n\left([a, b]\right)\right] = \frac{1}{\sqrt{3}} \nu_K([a, b]),$$

where ν_K is the equilibrium measure of K.

This is a special case of the following result, where μ does not need to be STU regular. The asymptotic lower bound requires very little of μ.

Theorem 2.3. *Let μ be a measure on the real line with compact support K.*
(a) Assume that $\mu' > 0$ a.e. in the interval $[a, b]$. Then

$$(2.6) \qquad\qquad \liminf_{n \to \infty} \frac{1}{n} \mathbb{E}\left[N_n\left([a, b]\right)\right] \geq \frac{1}{\sqrt{3}} \nu_K([a, b]).$$

(b) Suppose in addition that (2.4) holds, and that $[a, b] \subset O$. Then

$$(2.7) \qquad\qquad \limsup_{n \to \infty} \frac{1}{n} \mathbb{E}\left[N_n\left([a, b]\right)\right] \leq \frac{1}{\sqrt{3}} \inf_L \nu_L([a, b]),$$

where the inf is taken over all regular compact sets $L \subset K$ such that $L \supset [a, b]$, and the restriction $\mu|_L$ of μ to L is STU regular.

4

It is plausible that the right hand sides of (2.6) and (2.7) are equal under mild assumptions such as the one of part (a). An interesting open problem is to find rates of convergence in the limit relations (2.3) and (2.5).

3. PROOFS

Proof of Proposition 1.1. This proof is based on the discussions of Kac [17, p. 5-10] and Das [5]. The joint probability density of $\mathbf{c} = (c_0, c_1, \cdots, c_n)$ is

$$dP(\mathbf{c}) = (2\pi)^{-(n+1)/2} \sigma^{-(n+1)} e^{-\frac{\|\mathbf{c}\|^2}{2\sigma^2}} dc_0 dc_1 \cdots dc_n,$$

where $\|\mathbf{c}\|^2 = c_0^2 + c_1^2 + \cdots + c_n^2$. Since $G_n(x)$ has at most $M+1$ zeros in (a, b) for all $\mathbf{c}$ by Rolle's theorem, $N_n(a, b)$ is integrable over $\mathbb{R}^{n+1}$ with respect to $dP(\mathbf{c})$. Define

$$N_n^*(a, b) = N_n(a, b) - (\kappa(a) + \kappa(b))/2,$$

where

$$\kappa(x) = \begin{cases} 1 & \text{if } G_n(x) = 0, \\ 0 & \text{otherwise.} \end{cases}$$

Since $G_n(a)$ and $G_n(b)$ are continuous random variables, we have

$$\mathbb{E}[N_n(a, b)] = \int_{\mathbb{R}^{n+1}} N_n^*(a, b) \, dP(\mathbf{c}).$$

We state the following result from Kac [18, Theorem 1].

Lemma 3.1. *If $f(x)$ is continuous for $\alpha \le x \le \beta$ and continuously differentiable for $\alpha < x < \beta$, and $f'(x)$ vanishes only at a finite number of points in $\alpha < x < \beta$, then the number of zeros of $f(x)$ in $\alpha < x < \beta$ (multiple zeros are counted once and if either α or β is a zero, it is counted as $1/2$) is equal to*

$$\text{P.V.} \ \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{\alpha}^{\beta} \cos(yf(x)) \, |f'(x)| \, dx \, dy.$$

In our notation, this gives

$$N_n^*(a, b) = \text{P.V.} \ \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{a}^{b} \cos(yG_n(x)) \, |G_n'(x)| \, dx \, dy.$$

Thus

$$\mathbb{E}[N_n(a, b)] = (2\pi)^{-\frac{n+1}{2}} \sigma^{-(n+1)} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} N_n^*(a, b) e^{-\frac{\|\mathbf{c}\|^2}{2\sigma^2}} dc_0 dc_1 \cdots dc_n$$

$$(3.1) \qquad = \frac{\sigma^{-(n+1)}}{2\pi} \int_{a}^{b} \int_{-\infty}^{\infty} R_n(x, y) \, dy \, dx,$$

where

$$(3.2) \qquad R_n(x, y) = (2\pi)^{-\frac{n+1}{2}} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{-\frac{\|\mathbf{c}\|^2}{2\sigma^2}} \cos(yG_n(x)) \, |G_n'(x)| \, dc_0 dc_1 \cdots dc_n.$$

5

The interchange of the integration order is justified by the fact that the integrand is dominated by

$$e^{-\frac{\|\mathbf{c}\|^2}{2\sigma^2}} \sum_{j=0}^{n} |c_j| \, |g_j'(x)| \, ,$$

which is exponentially small outside bounded sets in $\mathbb{R}^{n+1}$. We use the known relation

$$(3.3) \qquad \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{1 - \cos(uv)}{u^2} \, du = |v|$$

to write (3.2) as

$$R_n(x, y) = \text{P.V.} \, \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{du}{u^2} \times$$

$$(3.4) \quad (2\pi)^{-\frac{n+1}{2}} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{-\frac{\|\mathbf{c}\|^2}{2\sigma^2}} \left(\cos(yG_n(x)) - \cos(yG_n(x))\cos(uG_n'(x)) \right) dc_0 dc_1 \cdots dc_n,$$

where the interchange of orders of the integration can be justified as above, and (3.4) is interpreted as

$$(3.5) \qquad \lim_{N \to \infty} \lim_{\epsilon \to 0} \frac{1}{\pi} \left(\int_{-N}^{-\epsilon} + \int_{\epsilon}^{N} \right) (\cdots) \frac{du}{u^2}.$$

Noting that

$$\cos(yG_n(x))\cos(uG_n'(x)) = \frac{1}{2}\Re\left(e^{iyG_n(x)+iuG_n'(x)} + e^{iyG_n(x)-iuG_n'(x)} \right),$$

we obtain with help of [13, 3.323(2) on p. 337] that

$$(2\pi)^{-\frac{n+1}{2}} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{-\frac{\|\mathbf{c}\|^2}{2\sigma^2}} \cos(yG_n(x))\cos(uG_n'(x)) \, dc_0 \cdots dc_n$$

$$= \frac{(2\pi)^{-\frac{n+1}{2}}}{2} \Re \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{-\frac{\|\mathbf{c}\|^2}{2\sigma^2}} \left(e^{i\sum_{j=0}^{n}[yc_j g_j(x)+uc_j g_j'(x)]} + e^{i\sum_{j=0}^{n}[yc_j g_j(x)-uc_j g_j'(x)]} \right) dc_0 \cdots dc_n$$

$$= \frac{(2\pi)^{-\frac{n+1}{2}}}{2} \Re \left(\prod_{j=0}^{n} \int_{-\infty}^{\infty} e^{-\frac{c_j^2}{2\sigma^2}+i[yg_j(x)+ug_j'(x)]c_j} \, dc_j + \prod_{j=0}^{n} \int_{-\infty}^{\infty} e^{-\frac{c_j^2}{2\sigma^2}+i[yg_j(x)-ug_j'(x)]c_j} \, dc_j \right)$$

$$= \frac{(2\pi)^{-\frac{n+1}{2}}}{2} \Re \left(\prod_{j=0}^{n} (2\pi)^{\frac{1}{2}}\sigma e^{-\frac{1}{2}[yg_j(x)+ug_j'(x)]^2\sigma^2} + \prod_{j=0}^{n} (2\pi)^{\frac{1}{2}}\sigma e^{-\frac{1}{2}[yg_j(x)-ug_j'(x)]^2\sigma^2} \right)$$

$$= \frac{\sigma^{n+1}}{2} e^{-\frac{\sigma^2}{2}\sum_{j=0}^{n}[yg_j(x)+ug_j'(x)]^2} + \frac{\sigma^{n+1}}{2} e^{-\frac{\sigma^2}{2}\sum_{j=0}^{n}[yg_j(x)-ug_j'(x)]^2}.$$

For $u = 0$, we have

$$(2\pi)^{-\frac{n+1}{2}} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{-\frac{\|\mathbf{c}\|^2}{2\sigma^2}} \cos(yG_n(x)) \, dc_0 \cdots dc_n = \sigma^{n+1} e^{-\frac{\sigma^2}{2}\sum_{j=0}^{n}[yg_j(x)]^2}.$$

Using abbreviations $A = A(x)$, $B = B(x)$ and $C = C(x)$, we rewrite

$$R_n(x, y) = \frac{\sigma^{n+1}}{\pi} \int_{-\infty}^{\infty} \frac{e^{-\frac{\sigma^2}{2}Ay^2}}{u^2} \, du$$
$$- \frac{\sigma^{n+1}}{2\pi} \left(\int_{-\infty}^{\infty} \frac{e^{-\frac{\sigma^2}{2}(Ay^2+Cu^2+2yuB)}}{u^2} \, du + \int_{-\infty}^{\infty} \frac{e^{-\frac{\sigma^2}{2}(Ay^2+Cu^2-2yuB)}}{u^2} \, du \right)$$
$$= \frac{\sigma^{n+1}}{\pi} e^{-\frac{\sigma^2}{2}Ay^2} \int_{-\infty}^{\infty} \frac{1 - e^{-\frac{\sigma^2}{2}Cu^2+yuB\sigma^2}}{u^2} \, du,$$

where the integral exists as a principal value, in the sense indicated in (3.5). If $C(x) = 0$ for some x then $B(x) = 0$ and $R(x, y) = 0$ for the same x and all y. Thus we set $By\sigma^2 = t$ and $\sigma^2 C = h > 0$, so that

$$R_n(x, y) = \frac{\sigma^{n+1}}{\pi} e^{-\frac{\sigma^2}{2}Ay^2} \int_{-\infty}^{\infty} \frac{1 - e^{-\frac{1}{2}hu^2+tu}}{u^2} \, du.$$

The Taylor expansion

$$e^{tu} = 1 + tu + \sum_{m=2}^{\infty} \frac{t^m u^m}{m!},$$

together with known identities

$$\int_{-\infty}^{\infty} \frac{1 - e^{-\frac{1}{2}hu^2}}{u^2} \, du = \sqrt{2\pi h} \quad \text{and} \quad \text{P.V.} \int_{-\infty}^{\infty} \frac{tu}{u^2} e^{-\frac{1}{2}hu^2} \, du = 0,$$

gives that

$$R_n(x, y) = \frac{\sigma^{n+1}}{\pi} e^{-\frac{\sigma^2}{2}Ay^2} \left(\sqrt{2\pi h} - \int_{-\infty}^{\infty} \left(\sum_{m=2}^{\infty} \frac{t^m u^m}{m!} \right) \frac{e^{-\frac{1}{2}hu^2}}{u^2} \, du \right).$$

Assuming that $h > 0$, we further obtain that

$$\int_{-\infty}^{\infty} \left(\sum_{m=2}^{\infty} \frac{t^m u^m}{m!} \right) \frac{e^{-\frac{1}{2}hu^2}}{u^2} \, du$$
$$= \sum_{m=1}^{\infty} \frac{t^{2m}}{(2m)!} \int_{-\infty}^{\infty} u^{2(m-1)} e^{-\frac{1}{2}hu^2} \, du$$
$$= \sum_{m=1}^{\infty} \frac{t^{2m}}{(2m)!} \frac{(2(m-1))!}{2^{m-1}(m-1)!} \sqrt{2\pi} h^{-(m-1)-\frac{1}{2}} \quad \text{(by [13, 3.461(2) on p. 364])}$$
$$= \sum_{m=1}^{\infty} \frac{\sqrt{2\pi h}}{m!(2m-1)} \left(\frac{t^2}{2h} \right)^m.$$

7

Hence

$$R_n(x,y) = \frac{\sigma^{n+1}}{\pi} e^{-\frac{\sigma^2}{2}Ay^2} \left(\sqrt{2\pi h} - \sum_{m=1}^{\infty} \frac{\sqrt{2\pi h}}{m!(2m-1)} \left(\frac{t^2}{2h} \right)^m \right)$$

$$= \sqrt{\frac{2C}{\pi}} \, \sigma^{n+2} e^{-\frac{\sigma^2}{2}Ay^2} \left(1 - \sum_{m=1}^{\infty} \frac{1}{m!(2m-1)} \left(\frac{B^2\sigma^2}{2C} \right)^m y^{2m} \right).$$

Applying [13, 3.461(2) on p. 364] again, we obtain that

$$\int_{-\infty}^{\infty} R_n(x,y)\,dy = \sqrt{\frac{2C}{\pi}} \, \sigma^{n+2} \int_{-\infty}^{\infty} e^{-\frac{\sigma^2}{2}Ay^2} \left(1 - \sum_{m=1}^{\infty} \frac{1}{m!(2m-1)} \left(\frac{B^2\sigma^2}{2C} \right)^m y^{2m} \right) dy$$

$$= \sqrt{\frac{2C}{\pi}} \, \sigma^{n+2} \left(\sqrt{\frac{2\pi}{A\sigma^2}} - \sum_{m=1}^{\infty} \frac{(\frac{B^2\sigma^2}{2C})^m}{m!(2m-1)} \frac{(2m)!}{m!2^m} \sqrt{\frac{2\pi}{A\sigma^2}} \frac{1}{(A\sigma^2)^m} \right)$$

$$= 2\sqrt{\frac{C}{A}} \, \sigma^{n+1} \left(-\sum_{m=0}^{\infty} \left(\frac{B^2}{AC} \right)^m \frac{(2m)!}{(m!)^2(2m-1)4^m} \right)$$

$$= 2\sqrt{\frac{C}{A}} \sqrt{1 - \frac{B^2}{AC}} \, \sigma^{n+1}.$$

Then (3.1) gives us the desired formula

$$\mathbb{E}[N_n(a,b)] = \frac{1}{\pi} \int_a^b \frac{\sqrt{AC - B^2}}{A} \, dx,$$

where $AC - B^2 \geq 0$ by (1.3) and the Cauchy-Schwarz inequality. $\qquad\square$

In addition to the reproducing kernels in (2.2), we also use their weighted versions in the proofs below:

$$\tilde{K}_n^{(k,\ell)}(x,y) = \mu'(x)^{1/2} \mu'(y)^{1/2} \sum_{j=0}^{n-1} p_j^{(k)}(x) p_j^{(\ell)}(y).$$

Lemma 3.2. *Let μ be a measure with compact support and with infinitely many points in its support. Let O be an open set in which μ is absolutely continuous, and such that for some $C > 1$ (2.4) holds. Then given any compact subinterval $[a,b]$ of O, we have*

$$(3.6) \qquad \frac{1}{n}\mathbb{E}\left[N_n([a,b])\right] = \frac{1+o(1)}{\sqrt{3}} \int_a^b \frac{1}{n} K_{n+1}(x,x)\,d\mu(x).$$

Proof. First note that the hypothesis that $\mu' \geq C^{-1}$ in O gives [12, Theorem 3.3, p. 104]

$$C_1 = \sup_{n\geq 1} \sup_{x\in[a,b]} \frac{1}{n} K_{n+1}(x,x) < \infty.$$

Next, we use Corollary 1.4 in [24, p. 224]. It gives for all $j, k \geq 0$,

$$(3.7) \qquad \lim_{n\to\infty} \int_a^b \left| \frac{\tilde{K}_{n+1}^{(j,k)}(x,x)}{\tilde{K}_{n+1}(x,x)^{j+k+1}} - \pi^{j+k}\tau_{j,k} \right| dx = 0.$$

8

Here

$$\tau_{j,k} = \begin{cases} 0, & j+k \text{ odd} \\ (-1)^{(j-k)/2}\frac{1}{j+k+1}, & j+k \text{ even}. \end{cases}$$

Applying (1.2) in a modified form, we obtain that

$$(3.8) \qquad \frac{1}{n}\mathbb{E}\left[N_n\left([a,b]\right)\right] = \frac{1}{\pi}\int_a^b \sqrt{\frac{\tilde{K}_{n+1}^{(1,1)}(x,x)}{\tilde{K}_{n+1}(x,x)^3} - \left(\frac{\tilde{K}_{n+1}^{(0,1)}(x,x)}{\tilde{K}_{n+1}(x,x)^2}\right)^2} \frac{1}{n}\tilde{K}_{n+1}(x,x)\,dx.$$

Since $\frac{1}{n}\tilde{K}_{n+1}(x,x)$ is bounded uniformly in n and in $x \in [a,b]$, we can use (3.7) above to obtain

$$\begin{aligned} \frac{1}{n}E\left[N_n([a,b])\right] &= \frac{1}{\pi}\int_a^b \left(\sqrt{\pi^2\tau_{1,1} - (\pi\tau_{0,1})^2} + o\left(1\right)\right)\frac{1}{n}\tilde{K}_{n+1}(x,x)\,dx \\ &= \frac{1+o(1)}{\sqrt{3}}\int_a^b \frac{1}{n}\tilde{K}_{n+1}(x,x)\,dx. \end{aligned}$$

$\square$

Proof of Theorem 2.2. Note that since $\mu' > 0$ a.e. in $[a,b]$, this interval is contained in supp ν_K. In [29, p. 287, Theorem 1], under weaker conditions, Totik proved that for a.e. $x \in [a,b]$,

$$\lim_{n\to\infty}\frac{1}{n}K_{n+1}(x,x) = \frac{d\nu_K}{d\mu}(x).$$

Since

$$\lim_{n\to\infty}\frac{1}{n}\tilde{K}_{n+1}(x,x) = \frac{d\nu_K}{d\mu}(x)\,\mu'(x) = \nu'_K(x),$$

the uniform boundedness of $\left\{\frac{1}{n}\tilde{K}_{n+1}(x,x)\right\}_{n=1}^{\infty}$ and Lemma 3.2 then give the result. $\square$

Proof of Theorem 2.3. We start with part (a). Given $r > 0$, and $j, k \geq 0$, with $\tau_{j,k}$ as above, it follows from [24, p. 250, Proof of Corollary 1.4] that

$$\left|\frac{\tilde{K}_{n+1}^{(j,k)}(x,x)}{\tilde{K}_{n+1}(x,x)^{j+k+1}} - \pi^{j+k}\tau_{j,k}\right|$$

$$\leq \frac{j!k!}{r^{j+k}}\sup_{|u|,|v|\leq r}\left|\frac{K_{n+1}\left(x+\frac{u}{\tilde{K}_{n+1}(x,x)}, x+\frac{v}{\tilde{K}_{n+1}(x,x)}\right)}{K_{n+1}(x,x)} - \frac{\sin\left(\pi\left(u-v\right)\right)}{\pi\left(u-v\right)}\right|.$$

Next, using that $\mu' > 0$ a.e. in $[a,b]$, we have from [24, p. 223, Theorem 1.1] that

$$\text{meas}\left\{x\in[a,b]:\sup_{|u|,|v|\leq r}\left|\frac{K_{n+1}\left(x+\frac{u}{\tilde{K}_{n+1}(x,x)}, x+\frac{v}{\tilde{K}_{n+1}(x,x)}\right)}{K_{n+1}(x,x)} - \frac{\sin(\pi(u-v))}{\pi(u-v)}\right| \geq \varepsilon\right\} \to 0$$

as $n \to \infty$, for any given $\varepsilon, r > 0$. Thus also

$$\text{meas}\left\{x\in[a,b]:\left|\frac{\tilde{K}_{n+1}^{(j,k)}(x,x)}{\tilde{K}_{n+1}(x,x)^{j+k+1}} - \pi^{j+k}\tau_{j,k}\right| \geq \varepsilon\right\} \to 0$$

9

as $n \to \infty$. Now let $\varepsilon > 0$, and for $n \geq 1$, let

$$\mathcal{E}_n = \left\{ x \in [a,b] : \sqrt{\frac{\tilde{K}_{n+1}^{(1,1)}(x,x)}{\tilde{K}_{n+1}(x,x)^3} - \left(\frac{\tilde{K}_{n+1}^{(0,1)}(x,x)}{\tilde{K}_{n+1}(x,x)^2}\right)^2} \leq \sqrt{\pi^2/3} - \varepsilon \right\}.$$

Then it follows that

$$\text{meas}\,(\mathcal{E}_n) \to 0 \text{ as } n \to \infty.$$

Using [30, p. 118, Thm. 2.1], we have for a.e. $x \in [a,b]$ that

$$\liminf_{n\to\infty} \frac{1}{n}\tilde{K}_{n+1}(x,x) \geq \nu_K'(x).$$

It then follows, that given $\varepsilon > 0$,

$$\mathcal{F}_n = \left\{ x \in [a,b] : \frac{1}{n}\tilde{K}_{n+1}(x,x) \leq \nu_K'(x) - \varepsilon \right\}$$

has

(3.9) $$\text{meas}\,(\mathcal{F}_n) \to 0 \text{ as } n \to \infty.$$

Indeed, if we set

$$f_n(x) = \min\left\{ \frac{1}{n}\tilde{K}_{n+1}(x,x) - \nu_K'(x), 0 \right\},$$

then by Totik's result,

$$\lim_{n\to\infty} f_n(x) = 0 \text{ a.e. in } [a,b],$$

while f_n is bounded below by $-\nu_K'$, so Lebesgue's Dominated Convergence Theorem gives

$$0 = \lim_{n\to\infty} \int_a^b f_n \leq \liminf_{n\to\infty}(-\varepsilon)\text{meas}\,(\mathcal{F}_n).$$

Thus (3.9) holds. Then by (1.2), (3.8) and the definitions of $\mathcal{E}_n$ and $\mathcal{F}_n$, we have

$$\frac{1}{n}\mathbb{E}\left[N_n\left([a,b]\right)\right] = \frac{1}{\pi}\int_a^b \sqrt{\frac{\tilde{K}_{n+1}^{(1,1)}(x,x)}{\tilde{K}_{n+1}(x,x)^3} - \left(\frac{\tilde{K}_{n+1}^{(0,1)}(x,x)}{\tilde{K}_{n+1}(x,x)^2}\right)^2}\,\frac{1}{n}\tilde{K}_{n+1}(x,x)\,dx$$

$$\geq \frac{1}{\pi}\int_{[a,b]\setminus(\mathcal{E}_n\cup\mathcal{F}_n)} \left(\sqrt{\pi^2/3} - \varepsilon\right)\left(\nu_K'(x) - \varepsilon\right)dx$$

$$\to \frac{1}{\pi}\int_a^b \left(\sqrt{\pi^2/3} - \varepsilon\right)\left(\nu_K'(x) - \varepsilon\right)dx \quad \text{as } n \to \infty.$$

Now we can let $\varepsilon \to 0$.

We pass to the proof of part (b). Let $L \subset K$ be a regular compact set such that the restriction $\mu|_L$ of μ to L is STU regular, and L contains $[a,b]$ in its interior. By monotonicity of the reproducing kernel (Christoffel function), if $K_n(\mu|_L, \cdot, \cdot)$ denotes the reproducing kernel of the measure $\mu|_L$, then for a.e. $x \in [a,b] \subset L$, Totik's result [29, p. 287, Theorem 1] gives

$$\limsup_{n\to\infty} \frac{1}{n}K_{n+1}(x,x)\mu'(x)$$

$$\leq \limsup_{n\to\infty} \frac{1}{n}K_{n+1}\left(\mu|_L, x, x\right)\mu'(x) = \nu_L'(x).$$

10

Moreover, $\left\{\frac{1}{n}K_{n+1}\left(\mu|_L, x, x\right)\mu'(x)\right\}_{n=1}^{\infty}$ is uniformly bounded in $[a, b]$. Then Lemma 3.2 implies that

$$\limsup_{n\to\infty}\frac{1}{n}\mathbb{E}\left[N_n\left([a, b]\right)\right] \leq \frac{1}{\sqrt{3}}\int_a^b \nu'_L(x)\, dx.$$

Finally, taking the inf over all L gives the result. $\qquad\square$

Lemma 3.3. *Let μ be an STU regular measure on the real line with compact support K, and let ν_K be the equilibrium measure of K. Suppose that the coefficients of random orthogonal polynomials (2.1) are complex i.i.d. random variables such that $\mathbb{E}[|\log|c_0||] < \infty$. If $E \subset \mathbb{C}$ is any compact set satisfying $\nu_K(\partial E) = 0$, then*

$$(3.10) \qquad \lim_{n\to\infty}\frac{1}{n}\mathbb{E}\left[N_n(E)\right] = \nu_K(E).$$

Proof. Consider the normalized counting measure $\tau_n = \frac{1}{n}\sum_{k=1}^{n}\delta_{z_k}$ for a polynomial (2.1), where $\{z_k\}_{k=1}^{n}$ are the zeros of that polynomial, and δ_z denotes the unit point mass at z. Theorem 2.2 of [25] implies that measures τ_n converge weakly to ν_K with probability one. Since $\nu_K(\partial E) = 0$, we obtain that $\tau_n|_E$ converges weakly to $\nu_K|_E$ with probability one by Theorem 0.5' of [20] and Theorem 2.1 of [3]. In particular, we have that the random variables $\tau_n(E) \to \nu_K(E)$ a.s. Hence this convergence holds in L^p sense by the Dominated Convergence Theorem, as $\tau_n(E)$ are uniformly bounded by 1, see Chapter 5 of [14]. It follows that

$$\lim_{n\to\infty}\mathbb{E}[|\tau_n(E) - \nu_K(E)|] = 0$$

for any compact set E such that $\nu_K(\partial E) = 0$, and

$$|\mathbb{E}[\tau_n(E) - \nu_K(E)]| \leq \mathbb{E}[|\tau_n(E) - \nu_K(E)|] \to 0 \quad \text{as } n \to \infty.$$

But $\mathbb{E}[\tau_n(E)] = \mathbb{E}[N_n(E)]/n$ and $\mathbb{E}[\nu_K(E)] = \nu_K(E)$, which immediately gives (3.10). $\quad\square$

Proof of Theorem 2.1. Given any $\varepsilon > 0$, we find a closed set S satisfying the assumptions, and obtain from Theorem 2.2 that

$$\lim_{n\to\infty}\frac{1}{n}\mathbb{E}\left[N_n\left([a, b]\right)\right] = \frac{1}{\sqrt{3}}\nu_K([a, b])$$

for any interval $[a, b] \subset K^\circ \setminus S$, where K° is the interior of K. Note that both $\mathbb{E}\left[N_n(H)\right]$ and $\nu_K(H)$ are additive functions of the set H. Moreover, they both vanish when H is a single point by (3.10), because ν_K is absolutely continuous with respect to Lebesgue measure on K, see [27, Lemma 4.4.1, p. 117]. Hence (3.10) gives that

$$\lim_{n\to\infty}\frac{1}{n}\mathbb{E}\left[N_n\left(\mathbb{R}\setminus S\right)\right] = \frac{1}{\sqrt{3}}\nu_K(\mathbb{R}\setminus S).$$

We can find finitely many open intervals $I_k \subset \mathbb{R}$, $k = 1, \ldots, m$, covering S, with total length $\sum_{k=1}^{m}|I_k| < 2\varepsilon$. Let $R_k = \{x + iy : x \in I_k,\ |y| < 1\}$, $k = 1, \ldots, m$, so that for $R = \cup_{k=1}^{m}R_k$ we have $S \subset R$ and $\nu_K(\partial R) = 0$. Applying Lemma 3.3 again, we obtain that

$$\limsup_{n\to\infty}\frac{1}{n}\mathbb{E}\left[N_n\left(S\right)\right] \leq \limsup_{n\to\infty}\frac{1}{n}\mathbb{E}\left[N_n\left(\overline{R}\right)\right] = \nu_K(\overline{R}\cap\mathbb{R}) = \nu_K\left(\cup_{k=1}^{m}\overline{I_k}\right),$$

Absolute continuity of ν_K with respect to dx implies that the last term in the above estimate tends to 0 as $\varepsilon \to 0$. Thus (2.3) follows. $\qquad\square$

11

References

1. A. T. Bharucha-Reid and M. Sambandham, *Random Polynomials*, Academic Press, Orlando, 1986.
2. A. Bloch and G. Pólya, *On the roots of certain algebraic equations*, Proc. London Math. Soc. **33** (1932), 102–114.
3. P. Billingsley, *Convergence of Probability Measures*, John Wiley & Sons, Inc., New York, 1999.
4. M. Das, *Real zeros of a random sum of orthogonal polynomials*, Proc. Amer. Math. Soc. **27** (1971), 147–153.
5. M. Das, *The average number of real zeros of a random trigonometric polynomial*, Proc. Camb. Phil. Soc. **64** (1968), 721–729.
6. M. Das and S. S. Bhatt, *Real roots of random harmonic equations*, Indian J. Pure Appl. Math. **13** (1982), 411–420.
7. A. Edelman and E. Kostlan, *How many zeros of a random polynomial are real?*, Bull. Amer. Math. Soc. **32** (1995), 1–37.
8. P. Erdős and A. C. Offord, *On the number of real roots of a random algebraic equation*, Proc. London Math. Soc. **6** (1956), 139–160.
9. K. Farahmand, *Topics in Random Polynomials*, Pitman Res. Notes Math. **393** (1998).
10. K. Farahmand, *Level crossings of a random orthogonal polynomial*, Analysis **16** (1996), 245–253.
11. K. Farahmand, *On random orthogonal polynomials*, J. Appl. Math. Stochastic Anal. **14** (2001), 265–274.
12. G. Freud, *Orthogonal Polynomials*, Akademiai Kiado/Pergamon Press, Budapest, 1971.
13. I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series and Products*, Academic Press, Burlington, 2007.
14. A. Gut, *Probability: A Graduate Course*, Springer, New York, 2005.
15. I. A. Ibragimov and N. B. Maslova, *The average number of zeros of random polynomials*, Vestnik Leningrad University **23** (1968), 171–172.
16. I. A. Ibragimov and N. B. Maslova, *The mean number of real zeros of random polynomials. I. Coefficients with zero mean*, Theory Probab. Appl. **16** (1971), 228–248.
17. M. Kac, *On the average number of real roots of a random algebraic equation*, Bull. Amer. Math. Soc. **49** (1943), 314–320.
18. M. Kac, *On the average number of real roots of a random algebraic equation. II*, Proc. London Math. Soc. **50** (1948), 390–408.
19. M. Kac, *Nature of probability reasoning*, Probability and related topics in physical sciences, Proceedings of the Summer Seminar, Boulder, Colo., 1957, Vol. I Interscience Publishers, London-New York, 1959.
20. N. S. Landkof, *Foundations of Modern Potential Theory*, Springer-Verlag, New York - Heidelberg, 1972.
21. J. E. Littlewood and A. C. Offord, *On the number of real roots of a random algebraic equation*, J. Lond. Math. Soc. **13** (1938), 288–295.
22. J. E. Littlewood and A. C. Offord, *On the number of real roots of a random algebraic equation. II*, Proc. Camb. Philos. Soc. **35** (1939), 133–148.
23. D. S. Lubinsky, *A new approach to universality limits involving orthogonal polynomials*, Ann. Math. **170** (2009), 915–939.
24. D.S. Lubinsky, *Bulk universality holds in measure for compactly supported measures*, J. Anal. Math. **116** (2012), 219–253.
25. I. E. Pritsker, *Zero distribution of random polynomials*, preprint. arXiv:1409.1631
26. T. Ransford, *Potential Theory in the Complex Plane*, Cambridge Univ. Press, Cambridge, 1995.
27. H. Stahl and V. Totik, *General Orthogonal Polynomials*, Cambridge Univ. Press, Cambridge, 1992.
28. D. C. Stevens, *The average number of real zeros of a random polynomial*, Comm. Pure Appl. Math. **22** (1969), 457–477.
29. V. Totik, *Asymptotics for Christoffel Functions for general measures on the real line*, J. Anal. Math. **81** (2000), 283–303.
30. V. Totik, *Asymptotics of Christoffel functions on arcs and curves*, Adv. Math. **252** (2014), 114–149.
31. N. Trefethen, *Roots of random polynomials on an interval*, CHEBFUN web page http://www.chebfun.org/examples/roots/RandomPolys.html

32. Y. J. Wang, *Bounds on the average number of real roots of a random algebraic equation*, Chinese Ann. Math. Ser. A. **4** (1983), 601–605.

33. J. E. Wilkins, Jr. *An asymptotic expansion for the expected number of real zeros of a random polynomial*, Proc. Amer. Math. Soc. **103** (1988), 1249–1258.

34. J. E. Wilkins, Jr. *The expected value of the number of real zeros of a random sum of Legendre polynomials*, Proc. Amer. Math. Soc. **125** (1997), 1531–1536.

School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332, USA
E-mail address: lubinsky@math.gatech.edu

Department of Mathematics, Oklahoma State University, Stilwater, OK 74078, USA
E-mail address: igor@math.okstate.edu

Department of Mathematics, Oklahoma State University, Stilwater, OK 74078, USA
E-mail address: sophia.xie@okstate.edu